Fully automated breast boundary and pectoral muscle segmentation in mammograms

Fully Automated Breast Boundary and Pectoral Muscle Segmentation in Mammograms.

Andrik Rampun, Philip J. Morrow, Bryan W. Scotney

School of Computing and Information Engineering, Ulster University, Coleraine, N. Ireland, BT52 1SA

John Winder

School of Health Sciences, Institute of Nursing and Health, Ulster University, Newtownabbey, N. Ireland, BT37 0QB

Abstract

Breast and pectoral muscle segmentation is an essential pre-processing step for the subsequent processes in Computer Aided Diagnosis (CAD) systems. Estimating the breast and pectoral boundaries is a difficult task especially in mammograms due to artifacts, homogeneity between the pectoral and breast regions, and low contrast along the skin-air boundary. In this paper, a breast boundary and pectoral muscle segmentation method in mammograms is proposed. For breast boundary estimation, we determine the initial breast boundary via thresholding and employ Active Contour Models without edges to search for the actual boundary. A post-processing technique is proposed to correct the overestimated boundary caused by artifacts. The pectoral muscle boundary is estimated using Canny edge detection and a pre-processing technique is proposed to remove noisy edges. Subsequently, we identify five edge features to find the edge that has the highest probability of being the initial pectoral contour and search for the actual boundary via contour growing. The segmentation results for the proposed method are compared with manual segmentations using

☆Fully documented templates are available in the elsarticle package on CTAN.

*Corresponding authors: A. Rampun, P. J. Morrow and B. W. Scotney

Email address: y.rampun, pj.morrow, bw.scotney@ulster.ac.uk (Andrik Rampun, Philip J. Morrow, Bryan W. Scotney)

322, 208 and 100 mammograms from the Mammographic Image Analysis Society (MIAS), INBreast and Breast Cancer Digital Repository (BCDR) databases, respectively. Experimental results show that the breast boundary and pectoral muscle estimation methods achieved dice similarity coefficients of 98.8% and 97.8% (MIAS), 98.9% and 89.6% (INBreast) and 99.2% and 91.9% (BCDR), respectively.

Keywords: Breast mammography, Breast Segmentation, Pectoral Segmentation, Computer Aided Diagnosis

1. Introduction

In 2013, more than 53,000 cases of invasive breast cancer diagnosed in the UK caused more than 11,000 deaths [1]. In the US, an estimated 246,660 new cases of invasive breast cancer are expected to be diagnosed in 2016 with more than 46,000 women expected to die [2]. These figures make breast cancer one of the most common types of cancer affecting women globally. Although the mortality rate in most developed countries has been decreasing since 2000 due to an increase in different types of screening methods, much effort still needs to be invested in fighting this disease.

Breast cancer screening mammography is a standard procedure to detect cancers at a very early stage. Unfortunately, it is impractical for radiologists to analyse hundreds of mammograms every day; the task is time consuming and exhausting, which leads to false positives or false negatives. The use of computer aided diagnosis (CAD) systems as a 'second reader opinion' is becoming popular due to its consistency, reliability and speed. In breast CAD, accurate breast segmentation is a crucial pre-processing step to speed up the subsequent processes without losing any important anatomical information. However, breast and pectoral muscle segmentation is a challenging task, especially in scanned mammograms, due to artifacts (e.g. duct tape, tags, light leakages and imperfections in the scanning process [3]), low contrast along the breast skin line, and homogeneity between pectoral and breast tissues. With the advent of advanced

machine learning methods such as Convolution Neural Networks (CNN) and deep learning, many authors are now using such methods for image segmentation and classification. However, in breast mammography building a robust network model based on intensity or textures can quite challenging. For example, finding a decision boundary between two classes is difficult when the appearance of the breast and pectoral muscle regions are very similar. In addition, these classifiers need a large amount of data (and ground truth) to build a reliable network. Unfortunately, obtaining ground truth data from a radiologist is difficult and time consuming. Other issues such as training time and complex parameters are also among further factors that make their use in mammography segmentation difficult.

In this paper, we propose a method for fully automatic breast and pectoral muscle segmentation. For breast segmentation, firstly we identify the initial breast boundary using a simple and robust thresholding technique. Subsequently, using the initial breast boundary we evolve the contour using the active contour (AC) method developed by Chan and Vese [4] on the entropy image feature rather than on the original image as performed in most studies (e.g. a study by Chen and Zwiggelaar [5]). For pectoral muscle segmentation, we developed a 2D breast model based on the breast segmentation results and identified five edge characteristics, namely length ($\hat{L}$), eccentricity (E_c), orientation (θ), intensity and extent (E_x). Using these edge features, we select the most appropriate edge as an initial pectoral boundary and 'grow' it based on the most similar intensity among its neighbouring pixels.

The novel contributions of our work are:

1. This is the first paper to have introduced the use of edge features/characteristics (eccentricity and extent) for pectoral muscle segmentation. To our knowledge, none of the existing methods have incorporated this information into their methods. Recent studies have only used length, intensity and orientation information.
2. We developed a 2D breast model that can be used for pre-processing,

post-processing and pectoral muscle segmentation. The model is robust in finding the global threshold value to automatically find an initial breast boundary that is close to the actual boundary.

From the contributions of our work, we identify four main advantages of the proposed method:

1. The proposed method is fully automatic and requires no interaction from the user. In contrast, the methods proposed by Chen and Zwiggelaar [5] and Ferrari *et.al* [6] require users to place 40 points along the mask boundary, which can be time-consuming to achieve accurately.
2. We use a simple and robust thresholding method combining the 2D breast model to get the initial breast boundary (initial seed points). Ferrari *et.al* [6] used the Llyod-Max binarisation procedure together with a chain-code method to find the initial breast boundary. Wirth and Stapinski [7] performed two-level threshold procedures and employed a least-squares best-fit piecewise quadratic curve to finalise the seed points identified in the previous step. In contrast, our method is straightforward and fast in finding an initial breast boundary that is close to the actual boundary.
3. The proposed method employes AC without edges [4], which is robust when dealing with breast boundaries with low contrast. To our knowledge, all the AC-based breast segmentation methods [5, 6, 7, 8, 9] employ typical AC models that rely on image gradient information. Where gradient information is absent along the breast boundary, AC models without edges are still able to find the boundary.
4. Unlike the method developed by Oliver *et.al* [32], the proposed method does not need training data, which eliminates the training phase (hence speeding up the whole process and reducing the computational complexity). Our 2D breast model is flexible regardless of the image size, whereas Oliver *et.al* [32] proposed a breast model that is applicable only to a specific image size depending on the dimensions of the images in the training set.

2. Literature Review

There are many methods that have been developed for breast boundary and pectoral muscle segmentation. However, only a few of them have been evaluated quantitatively using all of the images in the MIAS database [26], according to the recent review conducted by Mustra *et.al* [10] in 2016. For breast boundary segmentation, Czaplicka and Włodarczyk [11] used a combination of a global thresholding method based on minimising measures of fuzziness and applied Sobel edge detection to find the breast boundary. The method proposed by Mustra and Grgic [3] is also quite similar, except that they performed adaptive contrast enhancement on the original image before finding a threshold value. Another thresholding based method was proposed by Raba *et.al* [12], which used several threshold values to obtain overlapping masks. The final threshold value is the mean of the grey level located within the smallest and the largest masks. Masek and Attikiouzel [13] used a local thresholding method in conjunction with a minimum cross-entropy thresholding algorithm to find the breast boundary.

Alternative methods are gradient-based techniques which exploit the gradient information along the initial breast boundary. Zhou *et.al* [14] firstly extracted the background from the breast region by searching for the largest background peak from the grey level histogram of the image and performed a line-by-line gradient analysis from the top to the bottom of the image to find the breast boundary. In the last decade, AC methods (and their variants) have gained much attentions especially in the field of biomedical image analysis. Wirth and Stapinski [7] used a dual threshold approach to get the initial placement seed points and performed a least-squares best-fit of a piecewise quadratic curve to finalise the initial seeds. Subsequently, they employed the greedy AC of Williams and Shah [15] on the edge-enhanced mammogram to find the final breast boundary. The work by Chen and Zwiggelaar [5] initialised 40 points along the mask boundary which was obtained via thresholding using the minimum grey level value between peaks of the background and the breast tissue in the grey level histogram. Subsequently these points were evolved using a

contour growing technique. Their method is similar to the method proposed by Martí *et.al* [8], except that the method determines the seed points automatically by analysing first local maximum along the x axis at half of the height of the image. Ferrari *et.al* [6] compared the results of two variants of AC methods: the edge-based AC models [16] and an adaptive active deformable contour model [17]. Filter-based methods have also been applied to breast boundary and pectoral muscle segmentation. Casti *et.al* [18] used 18 Gabor filters to detect edges along the breast boundary on a transformed image and applied local thresholding and a false positive reduction method to capture edges belonging to the breast boundary. Subsequently, they performed edge linking to connect all edges along the breast boundary.

For pectoral muscle segmentation the study of Czaplicka and Włodarczyk [11] used an iterative multi-Otsu's thresholding to segment the pectoral region until specified conditions were met. The method of Mustra and Grgic [3] performed an image enhancement followed by thresholding to get the initial boundary of the pectoral muscle. Subsequently, they randomly selected 10 points for polynomial fitting of the muscle boundary. In a different approach developed by Chakraborty *et.al* [19], they first approximate the pectoral muscle boundary as a straight line based on features such as average gradient, position, and shape. The straight line is then tuned to a smooth curve which represents the pectoral boundary. Chen and Zwiggelaar [5] used a region-growing technique to remove the pectoral muscle, then manually placed a seed point on the border between the pectoral and breast regions. Subsequently, the seed was grown based on the similarity with the region's mean intensity. In contrast, the method proposed by Raba *et.al* [12] placed the initial seed point inside the pectoral muscle and used size restriction criteria to avoid over segmentation.

Ferrari *et.al* [20] proposed a more sophisticated method by using 48 Gabor filters to capture edges with orientations between 120° to 170°. In the next step, the authors computed the vector summation of the image magnitude and phase from each of the Gabor filter responses, followed by Sobel edge detection to estimate the initial pectoral boundary. Kwok *et.al* [21] used knowledge about

the position and shape of the pectoral muscle and estimated the pectoral edge using a straight line. The estimated straight line is then refined using iterative 'cliff detection' to estimate the actual boundary of the pectoral muscle. Another method is proposed by Karssemeijer [22], based on the Hough transform. The author first estimates the region of the pectoral muscle using a global threshold value. Subsequently, the gradient magnitude and direction of the pectoral region were calculated using the 3×3 Sobel operator. To detect the pectoral boundary a Hough transform is performed on the gradient image and the best line representing the actual boundary was selected based on criteria defined by the author.

For the existing methods discussed in this section, we identified several deficiencies. For breast boundary estimation, most thresholding methods take account of all grey levels in the image. The main drawback of this approach is that it does not take account of the non-homogeneity of the image background, which means that low contrast parts of the breast are considered as background, resulting in under segmentation [8]. By considering all grey levels in the image, the selection of the threshold value is influenced by artifacts (e.g. duct tape, tag, etc). On the other hand, edge-based AC models were applied on the original image to find the final boundary of the breast. Unfortunately, in many cases breast boundaries are unclear/obscured by noise in the original images. For pectoral muscle segmentation, straight-line estimation approaches are not able to detect pectoral boundaries that have curved shapes, whereas thresholding and region growing- techniques fail to work when the pectoral and breast regions are homogenous.

2.1. Active Contour Models

Since AC models are among the most popular methods used especially in medical image segmentation, there are many variations available in the literature. The main goal of these models is to evolve the initial object's boundary (determined via manual delineation or automatic thresholding) close to the actual object's boundary resulting a segmentation of an object. In this section we

summarise AC models into the following groups:

Edge-based models: Recently, Ciecholewski [43] developed an AC model using an inflation/deflation force with a damping coefficient function. Once the initial boundary of an object is found the inflation/deflation force for each of the nodes is damped using the damping coefficient function to estimate the actual boundary which enables the method to accurately estimate weak boundaries or edges in noisy regions. Another variant of edge-based models is the method proposed by Álvarez *et.al* [44] based on a morphological approach. The author claimed that faster and more stable models were achieved by combining the morphological operators associated with the Partial Differential Equation components. Ferrari *et.al* [6] introduced an adaptive active deformable contour model (AADCM) by making the balloon force adaptive depending on the image gradient (a higher value in homogeneous regions and slower near the breast boundary) and the external energy is computed based on the magnitude and direction of the image gradient.

Wirth *et.al* [7] employed an AC model based on a Greedy algorithm [15] which considers the energy from continuity, curvature and image gradient at each point as well as automatic placement of the initial contour. Both studies in [5, 8] used AC models based on the concept of attraction forces (which make the contour enter a region) and this can be achieved by manually selecting a set of seed points and grow them based on a combination of different criteria. Kass et.al [16] developed a model called "United Snakes" which aims to combine the best features of the various snake techniques while maintaining the simplicity and elegance of the original formulation to provide more flexibility and minimised user interaction which was later reported in [9]. Lobregt and Viergever [18] developed a discrete dynamic contour model which has a similar structure to the Geometrically Deformed Model (GDM) developed by Miller *et.al* [45, 46]. The developed model [18] firstly refines the estimated objects boundary defined by the user by minimizing the energy function followed by finding its actual boundary by taking account of the local contour curvature of the model and image features. For optimisation this process was repeated discretely until the

model reached an equilibrium state where velocity and acceleration are zero for each vertex.

Region-based models: One of the most popular AC region-based models is the method proposed by Chan and Vese [4] which is based on techniques of curve evolution, the Mumford-Shah functional for segmentation and level sets. Although the curve-evaluation applies the concept of energy minimisation like most of the existing AC models, the main difference with their method is that the stopping term is based on Mumford-Shah segmentation techniques instead of a gradient-based function which enables the model to detect objects with obscure boundaries or not defined by gradient. On the other hand, Li *et.al* [47] proposed a robust level set method that can deal with intensity inhomogeneities. Firstly, a local clustering criterion function is defined based on the local intensity clustering property. Secondly, this function is integrated with respect to the neighborhood center to give a global criterion of image segmentation which is defined as an energy in level set formulation that represents a region of an image. Finally, the image segmentation process can be achieved by minimising the energy in the level set function. Another AC region-based method which attempts to deal with intensity inhomogeneities was proposed by Zhang *et.al* [48]. Motivated from the local binary fitting (LBF) energy function, Zhang *et.al* introduced a local image fitting (LIF) energy function based on the local image information by minimising the difference between the fitted image and the original image. The LIF function is used to evolve the level set function and regularised it with a novel Gaussian kernel filtering after each iteration to enhance the smoothing capacity.

3. Methodology: Breast Boundary Estimation

Figure 1 shows an overview of our proposed breast segmentation approach. Firstly, we modelled the breast appearance in medio-lateral oblique (MLO) view in the mammogram image to identify the orientation of the breast. Secondly, artifacts were removed to avoid incorrect localisation of the initial breast bound-

ary. Noise reduction was applied using a combination of median and anisotropic diffusion filters [23, 24]. Using the 2D breast model we identify the best threshold value to get the initial breast mask (initial breast boundary). Subsequently, we fed the initial breast mask into the AC models using the entropy image rather than the original image and performed post-processing to detect over segmented boundaries. Other texture descriptors were investigated (e.g. features from the Grey Level Co-occurrence Matrix (GLCM)) but we chose entropy due to its simplicity and robustness to differentiate textures between the image background and the texture along the skin-air boundary. From now on, we denote the initial breast boundary (or initial seed points) as C_I and the final breast boundary as C_F.

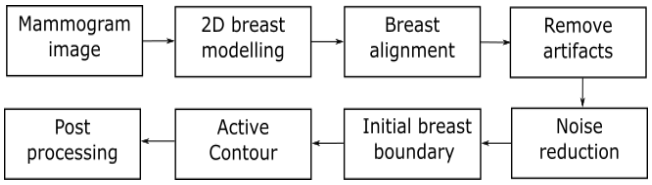

Figure 1: Overview of the breast boundary segmentation process

3.1. 2D Breast Modelling

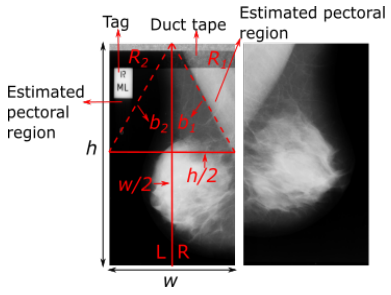

Figure 2: The initial 2D breast model for identification of breast orientation and threshold value in MLO view. Image *pdb002rl* (left) and *pdb001lm* (right) were taken from the MIAS database [26]

Figure 2 shows our breast 2D model for breast orientation detection. R_1 and R_2 represent the estimated right and left pectoral regions, respectively.

The pectoral boundary from each side is represented by the lines labelled b_1 (right) and b_2 (left) regardless to the location of the breast. These boundaries are the straight lines from the middle point of the image width (w) to half of the image height (h). The image is divided into left (L) and right (R) regions.

3.2. Breast Alignment

To standardise the workflow and the 2D breast model of the proposed method, if the breast is located in R (the right hand side of the image) we flip the image to the left by reflection in the vertical line at width $w/2$. Therefore, the proposed method can always assume that the location of the breast is within L. To detect the initial location of the breast we perform the following steps:

1. Apply Otsu's thresholding [25] on the original image.
2. Retain the largest region (the breast is always the largest region) from the resulting binary image. Keep all the other regions as an 'artifacts mask'.
3. Compute the entropy of the image obtained from the previous step using an $m \times k$ window size, where m and k are the number of rows and columns, respectively. In this study we used a 9×9 sliding window [27].
4. Sum the entropy in R_1 and R_2 and flip the image if it is greater in R_1 than in R_2.

Previous methods [5, 7, 8] simply compute average values in the left and right halves of the mammogram to decide the location of the breast. However this approach failed when large artifacts are present due to scanning imperfections (especially for images in the MIAS database).

3.3. Artifacts and Noise Removal

Since most scanned mammograms contain different artifacts, it is essential to remove them to minimise the risk of over segmentation. In the proposed method, this can be addressed by multiplying the 'artifacts mask' obtained from the previous stage with the original image. Subsequently, following the study of

Rampun *et.al* [27], we employed an Anisotropic Diffusion (AD) filtering [23, 24] to the median filtered image (9×9) to remove noise. Both filters worked in a complementary manner to remove different types of noise as shown in the study of Black *et.al* [28].

3.4. Estimating the Initial Breast Boundary

One of the disadvantages of using ACs in image segmentation is that the initial seed points (or initial mask) must be close to the actual or desired boundary. The most popular technique for finding C_I is via thresholding. A threshold value is determined by taking account of all grey levels in the entire image (e.g. methods in [5, 6, 7]). As mentioned in Section II, The main drawback of this approach is that it does not take account of the non-homogeneity of the image background, and so low contrast parts of the breast are considered as background. Hence the selection of the threshold value is influenced by artifacts, resulting in C_I being too far away from the actual boundary. To address this problem, we used our 2D breast model to find the possible grey levels belonging to the image background.

Since all images have been aligned so that the pectoral muscle is located at the top left of the image, we know that the majority of the image background is within the region R. In this phase, we are interested only in finding the range of the grey level values that belong to the image background. By multiplying the original image with the 'artifacts mask' image we are able to exclude a large number of grey levels belonging to the breast and artifacts. This technique minimises the risk of a threshold value being influenced by non-background grey levels (hence finding C_I that is close to the actual boundary). However, if no artifact is detected in the 'artifacts mask' then the process of multiplying the original image with the 'artifacts mask' will be skipped. This means that the artifact removal process is only used if duct tape or tag/label artifacts exist. To find the threshold value, we used equation 1

$$th = (med + sd)/2^b \tag{1}$$

where *med* and *sd* are the median and standard deviation of the grey levels in R, respectively, and b is the bit-depth of an image (e.g. 2^8 resulting in 256 grey levels for an 8-bit image). Note that at this stage each image has been normalised to the range [0,1]. Figure 3 shows four different breast masks using the proposed method, the histogram-based method of Chen and Zwiggelaar [5], the k-means thresholding of Mustra and Grgic [3], and Otsu's thresholding [25]. All threshold values in [3, 5, 25] were determined by taking account of all grey levels in the image, resulting in a poor (visually) separation between the breast and the background. In the present work, the initial seed points are all points along the red line in the left-most image in Figure 3. All coordinate points can be identified by employing Canny edge detection on the binary image and taking the longest connected edge.

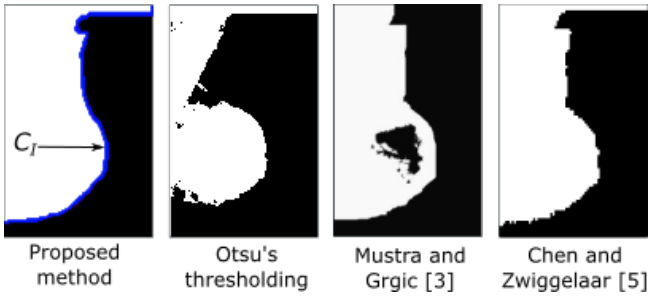

Figure 3: Qualitative comparison on different threshold values obtained from four different methods.

3.5. Estimating the Actual Breast Boundary

To our knowledge, all of the AC-based breast boundary estimation methods [5, 6, 7, 8] (and variants) used models with a stopping-edge function that relies on gradient information. The main disadvantage of these models is that they can detect only objects with edges defined by gradient [4]. In many cases edges are obscured along the skin-air breast boundary due to low contrast. In the proposed method, we employ AC models without edges, developed by Chan and Vese [4], that can accurately find the breast boundary along the skin-air.

The AC model of Chan and Vese [4] minimises the energy function $F(c_1, c_2, C)$, defined by

$$\begin{aligned} F(c_1, c_2, C_F) = \mu.\text{Length}(C_I) + \nu.\text{Area}(inside(C_I)) \\ +\lambda_1 \int_{inside(C_I)} |I(x,y) - c_1|^2 dx\ dy \\ +\lambda_2 \int_{outside(C_I)} |I(x,y) - c_2|^2 dx\ dy \end{aligned} \tag{2}$$

where c_1 and c_2 are the average intensity inside C (within the breast) and outside the C, respectively. The image is denoted as I, with each pixel located at coordinate (x,y). Chan and Vese [4] fixed the parameters as $\mu \geq 0$, $\nu = 0$, $\lambda_1 = \lambda_2 = 1$. Instead of using a stopping edge-function, the stopping term is based on Mumford-Shah segmentation techniques [29]. Using the breast mask, we employed AC models without edges [4] to search for C_F iteratively on the entropy image rather than on the original image as performed in [5, 6]. The entropy can be computed as

$$E = \sum_{j=1}^{J} p(j) \times \log_2(p(j)) \tag{3}$$

where J is the total number of grey levels, j is the j^{th} grey level and $p(j)$ is the probability of the j^{th} grey level obtained from the grey level histogram. We investigated more than 25 image features, including first and second order statistical features, and chose entropy due to its simplicity and ability to distinguish texture along the skin-air breast boundary. Entropy measures the randomness of the texture within a region of interest. Visually we can see that the textures within the breast are random (yielding a large number of grey levels) due to the appearance of different tissues such as fat, fibroglandular, breast and pectoral tissues, whereas the grey level variation within the background is limited. For the number of iterations (n) used to evolve C in the AC models, we set $n = 200$ for the entire experiment (see quantitative results in Figure 17). Once the segmentation is completed, we employed Canny edge detection on the binary image

and select the longest connected edge as the breast boundary. Other boundaries were considered as false positives and removed. Figure 4 shows examples of segmentation results when employing AC models on the entropy image and on the GLCM based features (e.g. sum of average and correlation) and original image (estimated breast boundary is shown in red on each image). The AC models failed to segment the skin-air region (under segmented) due to the low contrast in the GLCM based feature image and original image, but the contrast along the breast boundary increased in the entropy image and hence improved the segmentation results along the breast boundary significantly.

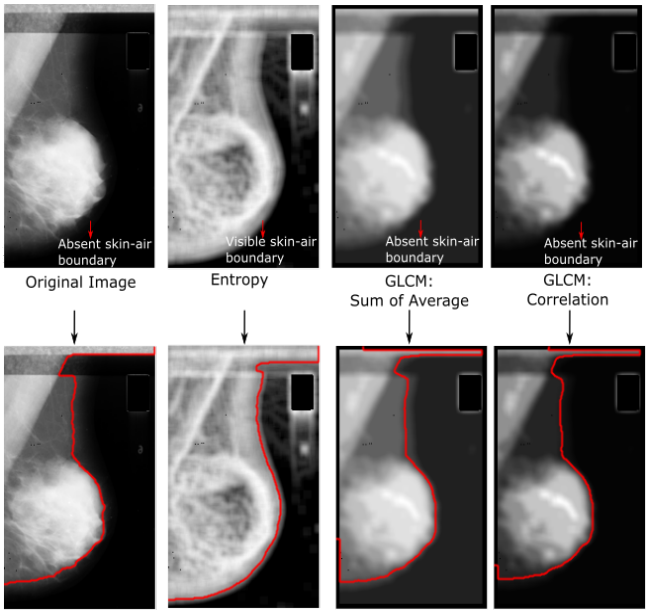

Figure 4: A visual comparison employing an AC on the computed entropy image and the original image. Note that the skin-air boundary is more visible in the entropy image.

3.6. Post-processing

In the proposed method, we use '*prior* knowledge' by identifying two critical regions in the image where artifacts are likely to occur. Several existing studies such as Mustra and Grgic [3], Kwok *et.al* [21] and Camilus *et.al* [30] have used '*prior* knowledge' based on their observations about the location, orientation

and appearance of the breast in mammograms. From our experience, we have found that the top right corner (G_2) and bottom left corner (G_1) as shown in Figure 5 usually suffer from over-segmentation due to presence of a labels/tags, duct tape or an imperfect scanning procedure. Therefore, we defined regions G_1 and G_2 where the breast boundary starting point (C_S) and end point (C_E) are located. Furthermore, duct tape and labels/tags are most likely to be located within these regions. We consider three aspects, namely corners, local minima and location in conjunction with the 2D breast model in Figure 5 to identify whether C_F is overestimated. Figure 5 illustrates our post-processing approach:

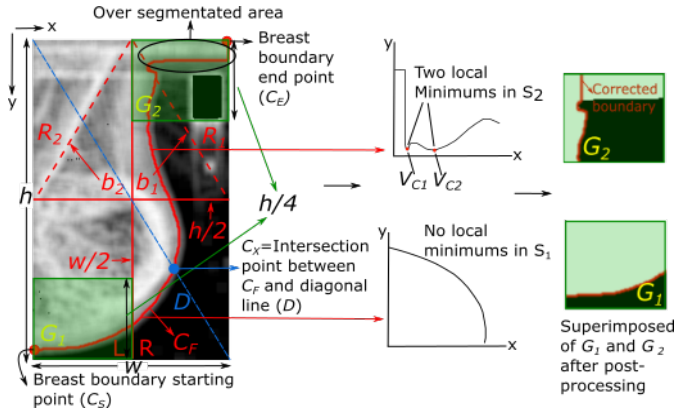

Figure 5: An overview of the post-processing approach. Note that the x and y axes in 2D space using the Cartesian coordinate system were taken from the y and x axes, respectively, from the image coordinates to capture local minimas.

1. Perform corner detection on the final segmented binary image using the corner detection method developed by Awrangjeb and Lu [31].
2. Treat C_F as a 1D signal consisting of two sub signals; S_1 is the set of all points from C_S to C_X and S_2 is the set of all points from C_X to C_E (see Figure 5). Smooth both sub-signals using convolution (a moving average is also possible) to remove false local minimas. Subsequently, plot each signal in 2D space using a Cartesian coordinate system and find all the local minimas in both signals (A local minimum is a data value that is smaller than its two neighbours).

3. Identify whether each corner is valid. A corner is valid (V_C) if it is located within a local neighbourhood of a local minimum (e.g. within 5×5). This approach removed false corners detected by the method developed by Awrangjeb and Lu [31].

4. If the location of the valid corner is within G_1 then we discard all points in S_1 for which $y > V_C(y)$ and connect V_C to the y axis (on the image), where $V_C(y)$ is the y coordinate of the valid corner. Similarly, if V_C is located in G_2, we discard all points in S_2 for which $x > V_C(x)$ and connect V_C to the x axis (on the image), where $V_C(x)$ is the x-coordinate of the valid corner. In the case where two V_Cs are detected in G_1 or G_2, we take the valid corner as the one that is closest to the y- and x- axis, respectively.

By the end of these steps, the estimated breast boundary is obtained.

4. Methodology: Pectoral Muscle Estimation

Figure 6 shows the workflow of our pectoral muscle segmentation method. We employed Canny edge detection on the smoothed image to detect initial

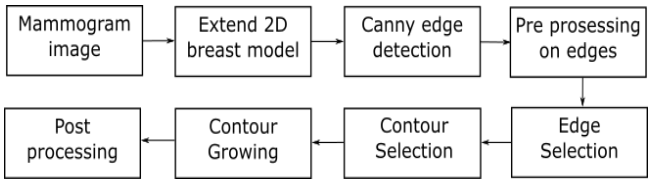

Figure 6: Overview of the pectoral muscle segmentation process.

contour candidates. Pre-processing was performed to ensure that the shapes and appearances of edges located along the actual pectoral muscle boundary are less affected by noise, artifacts and non-pectoral tissues. Edge selection is essential to remove false contours. Each selected edge was assumed to be a contour that represents a possible candidate for the pectoral boundary. We select the best edge based on its characteristics and 'grow' it using a contour growing technique to find the actual boundary. Finally we refine the boundary

using a robust local regression technique [35]. From now on, we use P_I and P_F to denote the initial and final pectoral boundary, respectively.

4.1. Extend 2D Breast Model

We extend the 2D breast model further by drawing a diagonal line (L) from C_E to C_S as a control boundary which will be used to determine the limit of the starting point (P_S) and end point (P_E) of the pectoral boundary. Figure 7 shows the final 2D breast model used in this study. Note that in this final model, C_E has been relocated after pre-processing.

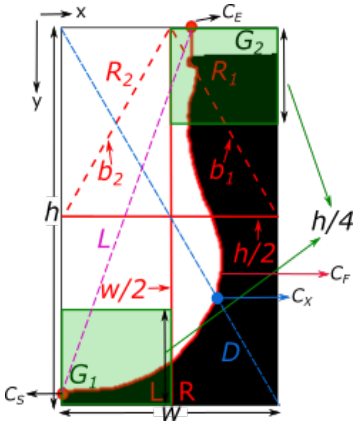

Figure 7: The final 2D breast model used in this study

4.2. Canny Edge Detection and Pre-processing

Canny edge detection [33] was employed to detect all possible P_I on the smoothed image and remove all edges below L (because the pectoral region is always located above L). Subsequently we 'clean' each edge to ensure that edges that belong to the pectoral boundary are disconnected from edges of other tissues such as fatty and fibroglandular tissue. The pre-processing phase in the proposed method removes connection points of the following edges using a 9×9 window. Figure 8 shows a 9×9 representation of the different edges.

1. Horizontal edges. We know that P_F is diagonal in orientation and most unlikely to be horizontal. The adjoining pixels ($\hat{C}(x \pm 1, y)$) of the central

pixel ($\hat{C}(x, y)$) in a $m \times k$ window are deleted if there are more than three connected pixels located in the same row to the left or right.

2. Branch edges. In some cases, the pectoral boundary is connected to other tissues that can alter the features of P_I. The diagonal adjoining pixels ($\hat{C}(x \pm 1, y \pm 1)$) are deleted if there are more than three connected pixels located in the top-left and bottom-right quadrant of the $m \times k$ window.
3. Half 'bullnose' edges. This is an edge for which the lower part is skewed to the right. We delete the connection points because the majority of P_F are either straight lines or a curve skewed to the left side of the image (the lower part of the pectoral boundary). The diagonal adjoining pixel ($\hat{C}(x + 1, y + 1)$) is deleted if there are more than three connected pixels located in the bottom-right quadrant of the $m \times k$ window.
4. Full 'bullnose' edges. This is an edge for which one of the points has the same y-coordinate value and discontinuous-x coordinate value with the other points in that edge. For example, the rightmost image in Figure 8 shows that the edge is categorised as full 'bullnose' because the value $y = 4$ appears three times in the fourth row and the x values from all points along this row are not continuous (e.g. $x = 5, 6, 9$). The diagonal adjoining pixels ($\hat{C}(x \pm 1, y \pm 1)$) are deleted if there are more than three connected pixels located in the quadrant to the bottom-right of the $m \times k$ window.

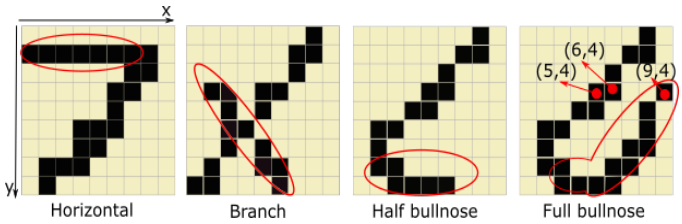

Figure 8: A 9×9 representation of the different edges which can alter the representation of the pectoral boundary.

Figure 9 shows an example of a binary image after edges below the L boundary were removed (middle) and the pre-processing step was applied (right most).

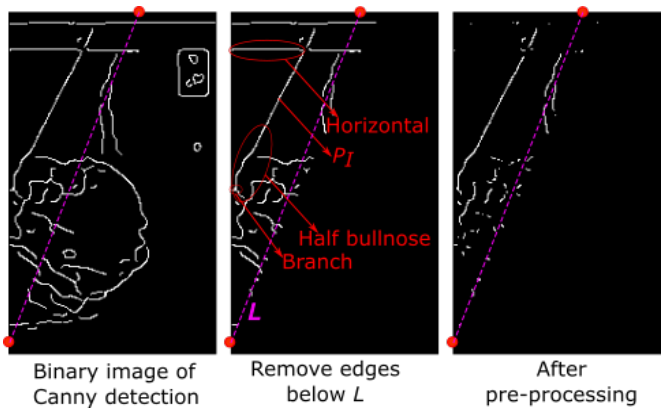

Figure 9: Example of binary images before and after pre-processing

4.3. Estimating the Initial Pectoral Boundary

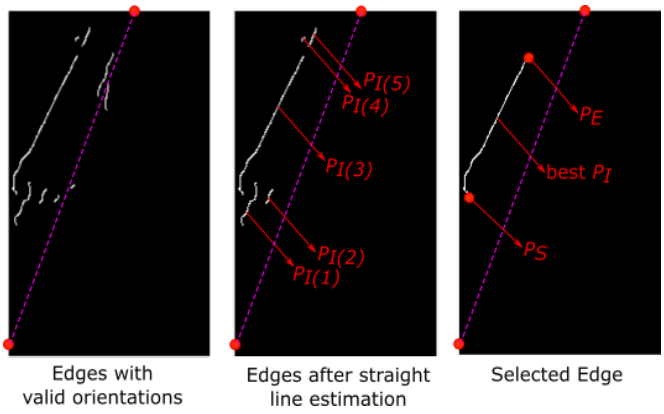

Figure 10: Resulting images representing each of the steps in selecting the best P_I.

Figure 10 shows the resulting binary images produced after each of the steps in this phase. In the first stage, we take only the remaining edges (right-most image in Figure 9) with θ between $30° - 90°$ (following the studies in [20, 22]). Subsequently, for each edge in left-most image (Figure 10), we estimate its straight line intersection at y and x axis. If its estimated straight line intercepts on the y- or x- axis is outside C_S and C_E, respectively, then we assume that

it does not belong to the pectoral boundary, because the pectoral boundary should not exceed the breast boundary. The straight line can be computed by taking the P_S and P_E coordinates of the edge to find the gradient (G) in the straight line equation $Y = GX + c$.

In the second stage we calculate features ($\hat{L}$, E_c and E_x) for each of the edges in the middle image in Figure 10. $\hat{L}$ represents the number of pixels of the edge and E_c calculates the ratio of the edge's length to the longest perpendicular chord of the ellipse [34]. The value of E_c is between 0 (a circle) and 1 (a line segment). The higher the value of E_c, the more probable that P_I is the actual pectoral contour. In addition, E_x calculates the ratio of edge's pixels to the total number of pixels in the bounding box [34]. A bounding box is a rectangle box which covers the two longitudes and two latitudes of the edge. The smaller the value E_x the more probable that P_I is the actual pectoral boundary. To select the best edge, we use the following steps:

1. Calculate $\hat{L}$ for each P_I and find the mean ($\hat{M}$) and standard deviation ($\hat{\sigma}$) for all P_I in the middle image in Figure 10. Find $\hat{T} = \hat{M} + \hat{\sigma}$.
2. If the number of edges $N(P_I)$ for which $\hat{L} > \hat{T}$ is one (there is only one edge that is significantly longer than the other edges), then the best edge is the longest P_I.
3. If $N(P_I)$ for which $\hat{L} > \hat{T}$ is more than one (there are a few edges which are significantly longer than the other edges), then we use a majority vote based on $\hat{L}$, E_c and E_x. Therefore, the best edge is the one with the highest vote. In case of a tie, the best edge is the longest P_I.
4. If $N(P_I)$ for which $\hat{L} > \hat{T}$ is zero (length of all edges are less than $\hat{M} \pm \hat{S}$), then we use a majority vote based on the same features ($\hat{L}$, E_c and E_x). Therefore, the best edge is the one with the highest vote. In case of a tie, the best edge is the longest P_I. In the middle image in Figure 10, $P_I(3)$ is selected as the best candidate.

If none of the above conditions are met then it is assumed that the pectoral muscle region does not appear in the image.

4.4. Estimating the Actual Pectoral Boundary and Pre-processing

To estimate the final pectoral boundary (P_F), we first calculate the average intensity value of P_I and connect P_E with the most similar intensity with the neighbour (located within the D_1 in Figure 11) of the closest seed in S_H. Similarly, this process is repeated for P_E and S_T but using the most similar intensity with the neighbour within the D_2 instead. S_T and S_H contains all the seeds along the straight lines from P_S to the y-intercept and P_E to the x-intercept, respectively. Finally, we used the Robust Local Regression MATLAB function to smooth the boundary.

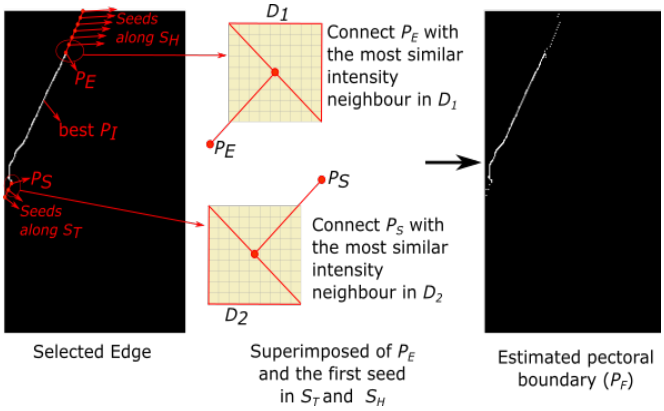

Figure 11: A graphical representation of growing best P_I to find P_F.

Figure 11 illustrates the final step of the proposed method. We connect P_E and P_S with the most similar intensity neighbour in D_1 and D_2, respectively, to ensure that the direction of the boundary is only upward at P_E and downward at P_S. This process is repeated until all seeds in S_H and S_T are covered. The pectoral contour is obtained by the end of this process.

5. Experimental Results

To test the performance of the proposed method, we used three different databases namely the MIAS database [26] (322 MLO mammograms), the INBreast database [41] (208 MLO mammograms) and the BCDR database [42]

(100 MLO mammograms) giving a total of 630 MLO mammograms. To reduce the required computational time, each image was resized by a factor of 4 similar to the studies in [3, 21, 32, 36]. Each scanned mammogram in the MIAS database contains manual segmentation (regarded here as ground truth, provided by the study of Oliver *et.al* [32]) for both breast and pectoral muscle boundaries annotated by a clinician supervised closely by an expert radiologist. For the INBreast database [41] the pectoral muscle boundary annotations are provided by an expert radiologist and the breast boundary annotations are provided by one of the authors. For the BCDR database [42], both pectoral and breast boundary annotations are provided by one of the authors. The method was developed under the MATLAB environment version 9 (2016a) on a Windows 10 operating system with an intel CORE i7 vPro processor.

5.1. Quantitative Results

To evaluate the performance of the proposed method we used the following metrics:

$$Jaccard\ (\ddot{J}) = \frac{|A \cap B|}{|A \cup B|} \tag{4}$$

$$Dice\ (\ddot{D}) = \frac{2\,|A \cap B|}{|A| + |B|} \tag{5}$$

$$Accuracy\ (\ddot{A}) = \frac{TP + TN}{TP + TN + FN + FP} \tag{6}$$

$$Sensitivity\ (\ddot{S}) = \frac{TP}{TP + FN} \tag{7}$$

$$Specificity\ (\bar{S}) = \frac{TN}{TN + FP} \tag{8}$$

$$Correctness\ (\ddot{C}) = \frac{TP}{TP + FP} \tag{9}$$

where TP, TN, FP and FN are true positive, true negative, false positive and false negative respectively. A is the number of common elements from segmented

region of the proposed method and B is the ground truth region. More details of these metrics can be found in [5, 6, 14, 19, 20, 27].

Table 1: Average Quantitative Results over 322 images from MIAS database. All metrics are presented as percentages with standard deviation ($\% \pm \sigma$).

Metric	(B_r+P_r) vs B_g	B_r vs (B_g+P_r)	P_r vs (B_g+B_r)	Mean
$\ddot{J}$	97.6 ± 1.9	95.1 ± 4.6	92.1 ± 13.6	94.9 ± 6.7
$\ddot{D}$	98.8 ± 1.1	97.3 ± 2.8	97.8 ± 11.8	97.9 ± 5.2
$\ddot{A}$	98.4 ± 1.3	97.1 ± 2.4	98.1 ± 1.8	97.9 ± 1.8
$\ddot{S}$	98.7 ± 1.9	97.5 ± 3.4	89.6 ± 12.5	95.3 ± 5.9
$\bar{S}$	97.6 ± 2.5	96.7 ± 3.1	98.9 ± 1.4	97.7 ± 2.3
$\ddot{C}$	98.8 ± 5.6	97.5 ± 6.3	95.6 ± 12.2	97.3 ± 8.0
Mean	98.3 ± 2.4	96.9 ± 3.8	95.3 ± 8.9	96.8 ± 4.9

Table 1 shows the overall quantitative results for the proposed method from MIAS database. We evaluate the performance based on three aspects: (a) breast and pectoral regions vs background ($(B_r + P_r)$ vs B_g) which evaluates the estimated breast boundary along the skin-air region; (b) breast region vs background and pectoral regions (B_r vs $(B_g + P_r)$), which measures the performance of the proposed method in separating the breast region only from the pectoral and air regions; and (c) pectoral region vs background and breast regions (P_r vs $(B_g + B_r)$), which indicates the performance of the proposed method in separating the pectoral region from the other regions. The quantitative results presented in Table 1 indicate that the proposed method produced very good results that are comparable with the best existing methods in the literature (Table 15). For breast boundary estimation, the proposed method achieved at least 97.6% in all metrics with the highest 98.8% on metrics $\ddot{D}$ and $\ddot{C}$. In terms of separating the breast region from the pectoral and air regions, the proposed method produced overlapping ratios of $\ddot{J} = 95.1\%$ and $\ddot{D} = 97.3\%$, which are close to $\ddot{S} = 97.5$. The results of the pectoral muscle boundary estima-

tion suggest that finding the pectoral boundary is more difficult than estimating the breast boundary, with $\hat{J} = 92.1\%$ and $\hat{S}$ close to 90%. A lower value for $\hat{S}$ indicates that the proposed method tends to produce under-segmented rather than over-segmented results. The standard deviation values for pectoral muscle segmentation are higher because there is a case (*pdb061ls* in Figure 12) for which the estimated boundary is far away from the actual pectoral boundary. The mean results for all metrics of $(B_r + P_r)$ vs B_g, B_r vs $(B_g + P_r)$, and P_r vs $(B_g + B_r)$ are 98.3%, 96.9% and 95.3%, respectively, which suggests that the proposed method is robust and comparable with existing methods in the literature.

Figure 12 shows example segmentation results from the MIAS database [26] together with its corresponding ground truth. The red and magenta lines represent the estimated pectoral and breast muscle boundaries for the proposed method, respectively. For each pair, the left-hand image is the ground truth for the right-hand image. We plot the estimated C_F (magenta) and P_F (red) on both ground truth and original images. The majority of the estimated breast and pectoral boundaries on images in the first row in Figure 12 achieved evaluation metrics of at least 88% for all metrics with a maximum of 99.9%. Moreover, as summarised in Table 1, the numerical evaluations for each case in Figure 12 further suggests that estimating the pectoral boundary is a difficult task. For example, for every case, the value of $\ddot{J}$ and $\ddot{D}$ for $(B_r + P_r)$ vs B_g are always higher than for P_r vs $(B_g + B_r)$. Images in the second row have been selected to show examples of inaccurate estimates for C_F or P_F. In the cases *pdb054rs* and *pdb062rs*, although C_F was estimated close to the ground truth, P_F was estimated incorrectly due to homogeneity between the breast and pectoral regions, which makes the actual pectoral boundary is difficult to discern. In the cases *pdb66rm* and *pdb61rl*, C_F was under-segmented due to homogeneity between the skin-air boundary and the air region. In the case *pdb065rl* the estimated boundary was far away from the actual pectoral boundary, resulting in the overall results producing higher σ as presented in Table 1.

On the other hand, Table 2 shows the overall quantitative results for the

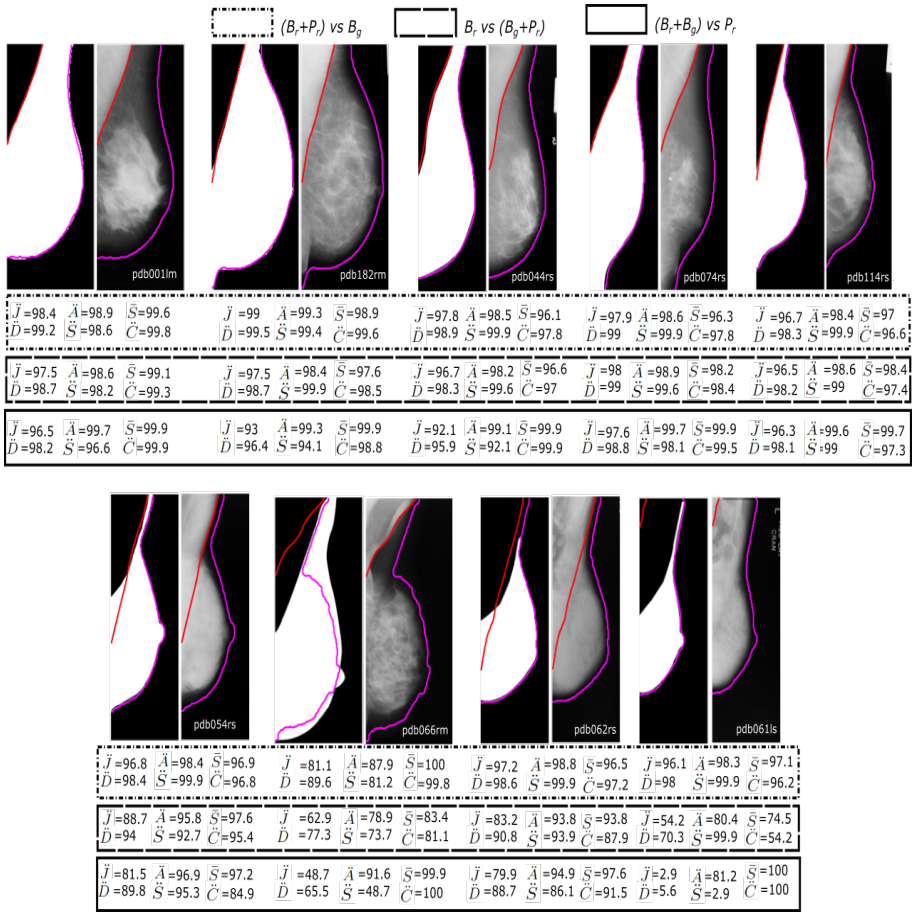

Figure 12: Examples of segmentation results. Images in the bottom row are cases where the proposed method failed to estimate the C_F or P_F accurately

proposed method using the INBreast database [41] which shows that the method produced very similar to those in Table 1. This indicates our method is robust in both breast and pectoral muscle segmentation. The proposed method produced
565 better results across different metrics in finding the breast boundary when tested on the INBreast database [41]. However, it can be observed that it produced at least 7% lower accuracy in finding the pectoral muscle boundary for metrics $\ddot{J}$ and $\ddot{D}$. On average, the proposed method produced $0.5 - 1.0\%$ better results across different metrics for $((B_r+P_r)$ vs $B_g)$ and $(B_r$ vs $(B_g+P_r))$ but 3.2%

Table 2: Average Quantitative Results over 208 images from INBreast database [41]. All metrics are presented as percentages with standard deviation ($\% \pm \sigma$).

Metric	(B_r+P_r) vs B_g	B_r vs (B_g+P_r)	P_r vs (B_g+B_r)	Mean
$\ddot{J}$	97.7 ± 0.77	95.9 ± 0.28	84.6 ± 15.6	92.7 ± 5.6
$\ddot{D}$	98.9 ± 0.41	97.4 ± 0.16	89.6 ± 10.1	95.3 ± 3.6
$\ddot{A}$	99.2 ± 0.27	98.3 ± 0.10	99.1 ± 0.9	98.9 ± 0.4
$\ddot{S}$	99.8 ± 0.52	98.9 ± 0.22	89.6 ± 9.6	96.1 ± 3.5
$\bar{S}$	98.7 ± 0.46	98.0 ± 0.11	99.7 ± 0.8	98.8 ± 0.5
$\ddot{C}$	98.9 ± 0.62	98.2 ± 0.21	90.2 ± 10.4	95.8 ± 3.8
Mean	98.9 ± 0.51	97.8 ± 0.18	92.1 ± 7.9	96.8 ± 4.9

worst in pectoral muscle segmentation (P_r vs (B_g+B_r)). Figure 13 shows several examples of breast (magenta line) and pectoral muscle (red line) segmentation from INBreast database [41]. Note that in the left most mammogram image in the second row, the proposed method successfully identified the pectoral muscle boundary despite the strong appearance of the *axillary fold*. The proposed method produced average results on the middle and right most mammogram images due to homogeneity between the pectoral muscle and breast region.

We further evaluated our method on 100 mammograms from the BCDR database [42]. Table 3 shows the overall quantitative results which are very similar to the results in Table 1 and Table 2. Once again this indicates that our proposed method is robust across different databases. However, as it can be observed, segmenting the pectoral muscle is a difficult task with $\ddot{J} = 85.8\% \pm 10.1$ and $\ddot{D} = 91.9\% \pm 6.8$ in comparison to finding the breast boundary ($\ddot{J} = 98.4\% \pm 0.4$ and $\ddot{D} = 99.2\% \pm 0.2$). Figure 14 shows several examples of breast (magenta line) and pectoral muscle (red line) segmentation from the BCDR database [42] which show all three cases were successfully segmented. Note that even the case where in the left most mammogram image with an obscure pectoral boundary the proposed method still managed to produce $\ddot{J} = 93.3\%$.

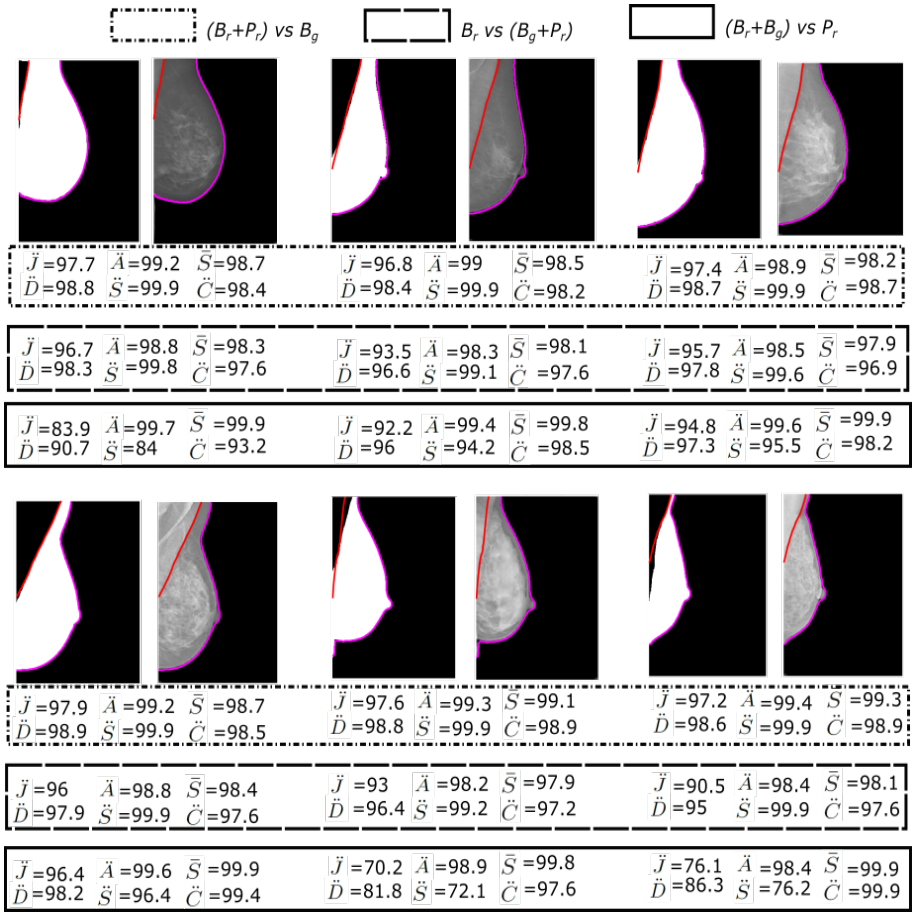

Figure 13: Examples of segmentation results from INBreast database [41].

To fully evaluate the performance of the proposed method, we compare our results with existing studies in the literature. It should be noted that it is difficult to make a direct comparison due to variations between datasets (e.g. number of images) and types of evaluation metrics (e.g. many studies used visual assessment by a radiologist to evaluate their methods). To minimise these variations, we made a comparison against studies that used the MIAS database [26] only and have similar metrics. The closest comparisons we can make are summarised in Table 4. Most of the methods in Table 4 did not test their algorithm on all 322 of the images in the MIAS database. Results for $(B_r + P_r)$

Table 3: Average Quantitative Results over 100 images from BCDR database [42]. All metrics are presented as percentages with standard deviation ($\% \pm \sigma$).

Metric	(B_r+P_r) vs B_g	B_r vs (B_g+P_r)	P_r vs (B_g+B_r)	Mean
$\ddot{J}$	98.4 ± 0.4	96.1 ± 1.6	85.8 ± 10.1	93.4 ± 4
$\ddot{D}$	99.2 ± 0.2	97.9 ± 0.8	91.9 ± 6.8	96.3 ± 2.6
$\ddot{A}$	99.9 ± 0.2	98.8 ± 0.6	99.4 ± 0.6	99.4 ± 0.5
$\ddot{S}$	99.2 ± 0.2	99.6 ± 0.6	92.7 ± 8.3	97.2 ± 3
$\bar{S}$	99.2 ± 0.2	98.4 ± 0.9	99.9 ± 0.2	99.2 ± 3.1
$\ddot{C}$	98.9 ± 0.3	98.7 ± 0.8	94.3 ± 9.3	97.3 ± 3.5
Mean	99.1 ± 0.3	98.3 ± 0.9	943 ± 5.9	97.1 ± 2.8

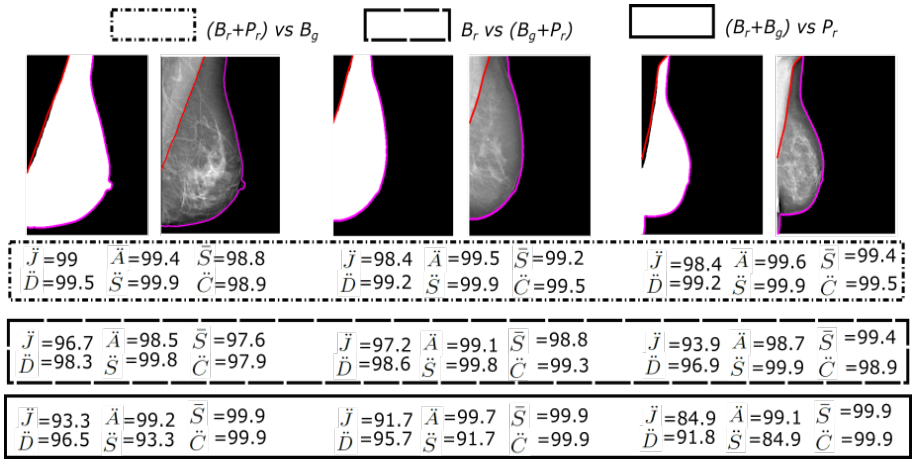

Figure 14: Examples of segmentation results from BCDR database [41].

vs B_g show that all methods achieved high accuracy of up to $\ddot{S} = 99.3\%$ and $\ddot{C} = 99.6\%$ from the study reported by Tzikopoulos *et.al* [36], which are slightly better than the results for our method ($\ddot{S} = 98.7\%$ and $\ddot{C} = 98.8\%$) and for the results in Wirth *et.al* [37] ($\ddot{S} = 99\%$ and $\ddot{C} = 98\%$). For a smaller dataset, Martí *et.al* [8] achieved $\ddot{S} = 96.9\%$ and $\ddot{C} = 95.5\%$. Oliver *et.al* [32], whose method used atlas, intensity and texture information in probability functions

Table 4: Qualitative Comparison. All results were quantitatively evaluated based on ground truth.

Authors	Dataset	Results
Wirth *et.al* [37]	120 images of MIAS	$(B_r + P_r)$ vs B_g: $\ddot{S} = 99\%$, $\ddot{C} = 98\%$
Tzikopoulos *et.al* [36]	All images in miniMIAS	$(B_r + P_r)$ vs B_g: $\ddot{S} = 99.3\%$, $\ddot{C} = 99.6\%$
Oliver *et.al* [32]	149 images of MIAS	$(B_r + P_r)$ vs B_g: $\ddot{D} = 96\%$ B_r vs $(B_g + P_r)$: $\ddot{D} = 97\%$ P_r vs $(B_g + B_r)$: $\ddot{D} = 83\%$
Martí *et.al* [8]	65 images of MIAS	$(B_r + P_r)$ vs B_g: $\ddot{S}$=96.9% , $\ddot{C}$=95.5%
Camilus *et.al* [38]	84 images of MIAS	P_r vs $(B_g + B_r)$: FP=0.85% and FN=4.88%
Ferrari *et.al* [20]	84 images of miniMIAS	P_r vs $(B_g + B_r)$: FP=0.58% and FN=5.77%
Liu *et.al* [39]	318 images of miniMIAS	P_r vs $(B_g + B_r)$: FP=3.34% and FN=4.57%

achieved $\ddot{D} = 96\%$, whereas our method achieved $\ddot{D} = 98.8\%$. Their method achieved a much lower value of $\ddot{D} = 83\%$ for P_r vs $(B_g + B_r)$, than our method (97.8%). Furthermore, performance for pectoral boundary estimation can be assessed using false positive and false negative rates among the methods in [20, 38, 39], which vary from less than 1% up to 6%. Liu *et.al* [39] achieved on average 2% FP rate on a larger number of images. All methods [20, 38, 39] produced very similar FN rates, ranging from 4.57% to 5.77%. Our method produced on average FP and FN rates of 0.52% and 2.75%, respectively over 322 images.

Figure 15 shows a visual comparison of the results produced by our method

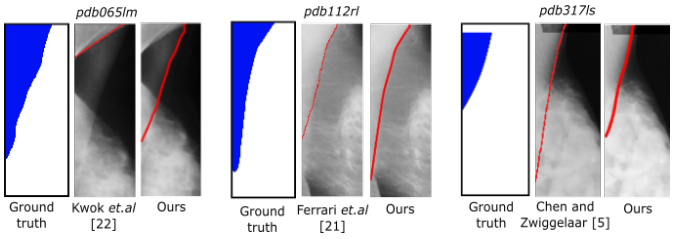

Figure 15: Visual comparison between our method against the methods of Kwok *et.al* [21], Ferrari *et.al* [20] and Chen and Zwiggelaar [5] from data of MIAS database. The blue region indicates the pectoral region. Note that the results were directly taken from the authors' publications for visual comparison purpose.

and the methods of Kwok *et.al* [21], Ferrari *et.al* [20] and Chen and Zwiggelaar [5] for three different cases. For case *pdb065lm*, the method of Kwok *et.al* [21] failed to estimate the pectoral boundary, resulting in the estimated contour being too far from the actual boundary, whereas our method is close with $\ddot{J} = 86.7\%$ and $\ddot{D} = 92.9\%$. Ferrari *et.al* [20] whose method used Gabor filers and edge flow propagation, underestimated the actual pectoral boundary in *pdb112rl*. In this case our method achieved $\ddot{J} = 91\%$ and $\ddot{D} = 95.3\%$. The method of Chen and Zwiggelaar [5] overestimated the pectoral boundary in image *pdb317ls* due to homogeneity between the pectoral and breast regions. In this case our proposed method achieved $\ddot{J} = 78\%$ and $\ddot{D} = 87.6\%$.

With respect to computationally complexity (time efficiency measured in seconds (s)), Table 5 shows the efficiency of the proposed method at different stages. The proposed method took on average under $50s$ to segment each image in the BCDR and INBreast databases (original images were resized by a factor of four), whereas images from the MIAS database took on average under $7s$. The AC models used in this study took on average over $30s$ to estimate the breast boundary for larger image dimensions compared to $3.5s$ for smaller image sizes (MIAS). It can be clearly observed that pectoral muscle segmentation is much faster compared to the breast boundary segmentation due to the inefficiency of the AC models. The overall time taken to segment all images in MIAS, BCDR and INBreast is approximately 37 minutes, 56 minutes and 2 hours 30 minutes,

respectively.

For computationally complexity (time efficiency measured in seconds (s)), Table 5 shows the efficiency of the proposed method at different stages. The proposed method took on average under $50s$ to segment each image in BCDR and INBreast databases (original images were resized by factor four), where images from MIAS database took on average under $7s$. The AC models used in this study took on average over $30s$ to estimate the breast boundary in a larger image dimension compared to $3.5s$ in a smaller image size (MIAS).It can be clearly observed that pectoral muscle segmentation is much faster compared to the breast boundary segmentation due to the inefficiency of the AC models. The overall time taken to segment all images in MIAS, BCDR and INBreast are approximately 37 minutes, 56 minutes and 2 hours 30 minutes, respectively. The proposed method took a shorter time to segment all the images in MIAS due to the small sizes of the images compared to BCDR and INBreast databases. On the other hand, the proposed method took a longer time to segment each image in the INBreast database because the initial breast boundary is much further away from the actual boundary (hence it took longer time to reach the actual boundary) and the number of possible candidates (the number of edges which have similar pectoral characteristics) is more.

Table 5: Experiment results for time efficiency at different stages of the proposed method. Note that all are measured in seconds (s).

Database	$\#_{img}$	O_{Dim}	T_{Dim}	μ_{ACWE}	μ_{BB}	μ_{PB}	t_{img}	$\sum t$
MIAS [26]	322	Smallest: 1080×400 Largest: 1300×1000	Smallest: 270×100 Largest: 325×250	3.5 ± 0.68	5.7 ± 7.7	1.2 ± 0.3	6.9 ± 7.9	1350
BCDR [42]	100	3328×4048	832×1012	31.8 ± 19.9	34.1 ± 24.5	15.2 ± 5.3	49.3×21.3	3373
INBreast [41]	208	4048×3328	1012×832	31.2 ± 24.3	33.3 ± 20.7	9.9 ± 4.2	43.2 ± 20.6	8865

$\#_{img}$: The total number of images in the database.

O_{Dim} : The original dimension of the image.

T_{Dim} : The new dimension of the image after downsized by factor 4.

μ_{ACWE} : Average time taken by the contour evolving from the initial breast boundary to the actual breast boundary.

μ_{BB} : Average time taken to segment the breast boundary including the time taken for pre-processing, estimating the initial breast boundary and post-processing.

μ_{PB} : Average time taken to segment the pectoral muscle boundary including pre-processing and post processing.

t_{img} : Average time taken to segment both breast and pectoral muscle boundaries.

$\sum t$: Overall time taken to segment all images in the database.

In addition, the computational complexity of inference of the proposed method is much simpler compared to those based on machine learning techniques (e.g. deep learning and convolution neural network (CNN)). Firstly, the complexity of inference for the proposed method is not influenced by the complexity of the predictive models built during the training process. Secondly, the parameters of the predictive models in CNN are more sensitive due to the large ranges of possible values (e.g. number of layers or neurons) whereas our method is less sensitive because we know roughly the range of orientation and area of the pectoral muscle boundary. Thirdly, the complexity of inference in machine learning based methods is more complex because it takes account of every single pixel and its neighbourhood. In contrast, our method only takes account of the most prominent pixels such as edge pixels. Thirdly, in most machine learning based methods, many features are required to build more accurate results whereas our methods need only take account of orientation, intensity, length, eccentricity and extent which makes the decision work flow simpler. Finally, when estimating the pectoral muscle boundary the number of pixels to be considered is smaller because only pixels around the edges are considered whereas in machine learning based methods, all pixels in the image will be considered which makes the computational complexity of inference is more complex.

6. Discussion

In terms of separating $(B_r + P_r)$ vs B_g, the proposed method is robust due to our initial seeds being close to the actual breast boundary. This was achieved by finding a threshold value from R instead of from the whole image, resulting in our threshold value being less influenced by the breast and pectoral muscle or by artifacts. Also our proposed method used AC models [4] that do not rely on gradient information but are based on Mumford-Shah segmentation techniques, making them robust in finding breast boundaries which cannot be represented well by image gradient, especially along the skin-air interface.

When separating P_r vs $(B_r + B_g)$, the proposed method is effective in finding

the actual pectoral muscle through the identification of edge features, namely orientation (θ), length ($\hat{L}$), eccentricity (E_c) and extent (E_x). For example, in a case where the *axillary fold* appears with high-intensity (Figure 15 *pdb065lm*), similar to the pectoral muscle, the method of Kwok *et.al* [21] failed to estimate the actual boundary because their method assumes that the initial pectoral muscle boundary is the one with the strongest gradient. Our method does not make such an assumption, with all edges treated as possible candidates, and the correct pectoral boundary is the one that satisfies the characteristic conditions.

However, there are three limitations of the proposed method. Firstly, since the pectoral muscle estimation relies on edges, in an event where the Canny method failed to detect any edges above the L region, the algorithm assumes that no pectoral muscle appears in the image. Secondly, selecting the correct P_I as an initial contour is crucial. If the initial contour selected is incorrect, then the final estimated boundary will also be wrong. Finally, the AC models used in this study could be quite slow due to the need to periodically reinitialise the model to repair the level set function degraded while the contour evolves during subsequent iterations.To overcome these problems, for future work we plan to use Gabor filters in conjunction with the Canny method for edge detection in order to increase the sensitivity of the method in detecting actual edges and to address the third issue we plan to employ the AC models developed by Li et.al [47] or Zhang et.al [48] which are claimed to be more efficient and robust by the authors.

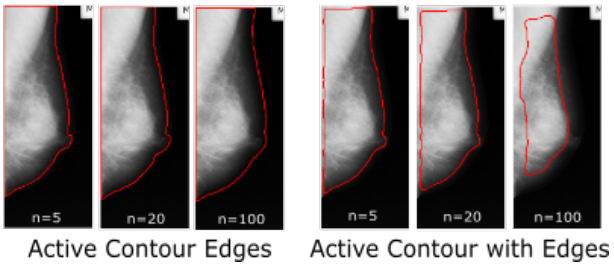

Figure 16: A visual comparison using AC without edges and AC with edges using different number of iterations.

For the selection of AC models [4], Figure 16 shows examples for visual comparisons between ACs without and with edge/gradient [40] information. We tested both methods for three different numbers of iterations from $n = 5$ to $n = 100$. Visually, it can be seen that AC models without edges [4] outperformed AC models with edges [40] regardless of the number of iterations. At $n = 5$, AC with edges [40] already under-segmented the breast boundary and failed to find the correct boundary at $n = 100$. This is because the contour is attracted to regions, with higher gradient which are usually located within the breast region due to the appearance of fibroglandular tissues, hence resulting in under-segmentation. In contrast, AC models without edges [4] are not attracted to edges/gradient and therefore tend to move towards the breast boundary and stop close to or on the skin-air boundary. This makes AC models without edges more suitable for our problem domain because edges are often difficult to discern along the skin-air boundary.

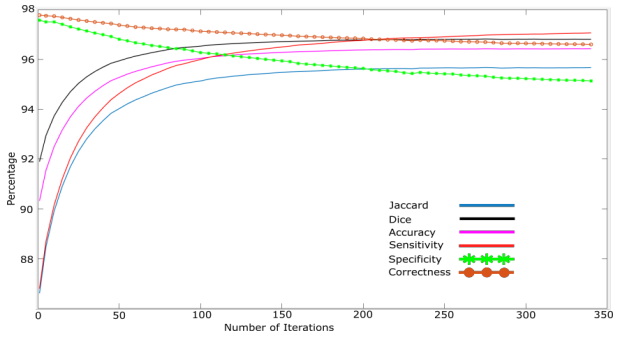

Figure 17: Experimental results on $B_r + P_r$ vs B_g using different number of iterations in the AC models on MIAS database.

To investigate the robustness of the proposed method to the number of iteration(n), we evaluated the results for $(B_r + P_r)$ vs B_g using 68 different iterations ranging from $n = 1$ to $n = 340$ (at intervals of five iterations n)on MIAS database. The results shown in Figure 17 indicate that most evaluation metrics achieved their best values after $n \geq 200$ (no further change after 200 iterations). This indicates that the P_F and C_F were estimated very close to

the actual boundaries after 200 iterations. Most evaluation metrics increased rapidly between $n = 1$ and $n = 50$, which suggests that the minimum number of iterations is 50. From $n = 51$ to $n = 199$ there is a small improvement of between 2% and 3.5% for metrics $\ddot{J}$, $\ddot{D}$, $\ddot{A}$ and $\ddot{S}$. On the other hand, the trend for metrics $\bar{S}$ and $\ddot{J}$ is a decrease between $n = 1$ and 340 due to an increase in the number of false positives. Note that, based on the best n value here, we used the same number of iterations for both INBreast [41] and BDCR databases (results shown in Tables 2 and 3, respectively).

In terms of the parameter settings in the AC models used in this study, we set the smoothing factor $\mu = 0.02$. In our experiments across three different databases, we found that this parameter did not give much variability on the segmentation results due to the following: (a) most of the breast boundary in a mammogram is usually smooth hence increasing or decreasing the value of μ does not have much effect on the segmentation results, (b) the entropy image feature enhances the contrast along the skin-air boundary which makes it easier for the AC models to estimate the actual breast boundary and (c) in a case where a jagged boundary is estimated, this will be automatically smoothed in the post-processing using convolution. In addition, for the time step (δt) and space step (h) we set the values to 0.1 and 1.0, respectively as used in the study of Chan and Vese [4].

In terms of the selection of window size (ws) we have performed several experiments using different values for ws. Table 6 presents the performance of the proposed method when different values of ws are used to estimate the breast boundary for INBreast database [41]. Generally, using the smallest value of ws produces the best results across different metrics. However, the variation in the results is very small which indicates that the proposed method produces consistent results across different window sizes. In fact, the average performance shows that our method achieved results similar to the state-of-the-art in the literature.

Table 6: Experimental results across different window sizes for $(B_r + P_r)$ vs B_g for INBreast database [41]. All measured in percentage (%).

Window size	$\ddot{J}$	$\ddot{D}$	$\ddot{A}$	$\ddot{S}$	$\bar{S}$	$\ddot{C}$
3×3	98.9	99.4	99.6	99.9	99.5	99.3
5×5	98.7	99.3	99.5	99.9	99.3	99.2
7×7	98.2	99.1	99.3	99.8	99.1	99.1
9×9	97.7	98.9	99.3	99.8	98.7	98.9
11×11	97.3	98.7	99.1	99.7	98.5	98.7
Mean	98.2	99.1	99.4	99.8	98.4	99

7. Summary and Conclusions

We have developed a new method for automatic segmentation of the breast boundary and pectoral muscle in MLO views of mammograms. In the breast boundary estimation, the method determines initial seeds that are close to the actual boundary and subsequently employs AC models that do not rely on gradient information, yielding accurate estimates for the boundary along the skin-air interface. For the pectoral boundary estimation, we used Canny edge detection to find all possible candidates and used edge features to determine the most probable pectoral contour followed by contour growing technique based on intensity to find the actual boundary. Moreover, we developed a 2D breast model that can be used for breast alignment to determine threshold values for the initial seeds, post-processing and false positive reduction for pectoral muscle estimation. The method was tested on three different databases (MIAS [26], INBreast [41] and BCDR [42]) which covered 630 MLO mamograms and evaluated based on ground truth from manual segmentation. Results for both breast and pectoral boundary segmentation show that the proposed method achieved high accuracy over a range of evaluation metrics and is comparable in performance to the best existing methods in the literature.

Acknowledgment

This research was undertaken as part of the Decision Support and Information Management System for Breast Cancer (DESIREE) project. The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 690238.

References

[1] Cancer Research UK, 'Breast cancer statistics', 2013. [Online]. Available: `http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer`. [Accessed:10-Oct-2016]

[2] BreastCancer.org, 'U.S. Breast Cancer Statistics', 2016. [Online]. Available: `http://www.breastcancer.org/symptoms/understand_bc/statistics`. [Accessed:10-Oct-2016]

[3] M. Mustra and M. Grgic. "Robust automatic breast and pectoral muscle segmentation from scanned mammograms", *Signal Processing*, vol. 93, no. 10, pp. 2817–282, 2013.

[4] T. F. Chan and L. A. Vese. "Active contour without edges", *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, 2001.

[5] C. Chen and Z. Zwiggelaar. "Segmentation of the breast region with pectoral muscle removal in mammograms", *in Proc. 14th Annu. Conf. on Medical Image Understanding and Analysis (MIUA)* 2010, pp. 71–75.

[6] R. J. Ferrari, F. Frère, R. M. Rangayyan, J. E. L. Desautels and R. A. Borges. "Identification of the breast boundary in mammograms using active contour models", *Med. Biol. Eng. Comput.*, vol. 42, no. 2, pp. 201–208, 2004.

[7] M. A. Wirth and A. Stapinski. "Segmentation of the breast region in mammograms using snakes", *in Proc. 1st Cannadian Conf. on Computer and Robot Vision (CRV)* 2004, pp. 385–392.

[8] R. Martí, A. Oliver, D. Raba and J. Freixenet. "Breast skin-line segmentation using contour growing", *in Proc. 3rd Iberian Conference, (IbPRIA)* (ser. Lecture Notes in Computer Science), vol. 4478, Springer Berlin Heidelberg, 2007, pp. 564–571.

[9] T. Ojala, J. Liang, J. Nappi and O. Nevalainen. "Interactive segmentation of the breast region from digitized mammograms with united snakes", Technical Report. Turku Centre for Computer Science, 1999.

[10] M. Mustra, M. Grgic and R. M. Rangayyan. "Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms", *Med. Biol. Eng. Comput.*, vol. 54, no. 7, pp. 1003–1024, 2016.

[11] K. Czaplicka and J. Włodarczyk. "Automatic breast-line and pectoral muscle segmentation", *Schedae Informaticae*, vol. 20, pp. 195–209, 2011.

[12] D. Raba, A. Oliver, J. Martí, M. Peracaula and J. Espunya. "Breast segmentation with pectoral muscle suppression on digital mammograms", *in Proc. 2nd Iberian Conference, (IbPRIA)* (ser. Lecture Notes in Computer Science), vol. 3523, Springer Berlin Heidelberg, 2005, pp. 471–478.

[13] M. Masek, Y. Attikiouzel and C. J. S. deSilva. "Skin-air interface extraction from mammograms using an automatic local thresholding algorithm", *in Proc. 15th Biennial International Conference Biosignal* 2000, pp. 204–206.

[14] C. Zhou, H. P. Chan, N. Petrick, M. A. Helvie, M. M. Goodsitt, B. Sahiner and L. M. Hadjiiski. "Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms", *Med. Phys.*, vol. 28, no. 6, pp. 1056–69, 2001.

[15] D. J. Williams and M. Shah. "A fast algorithm for active contours and curvature estimation", *CVGIP: Image Understanding*, vol. 55, no. 7, pp. 14–26 , 1992.

[16] M. Kass, A. Witkin and D. Terzopoulos. "Snakes: Active contour models", *Int. J. Comput. Vision*, vol. 55, no. 7, pp. 14–26 , 1992.

[17] S. Lobregt and A. Viergever. "A Discrete Dvnamic Contour Model ", *IEEE Trans. Med. Imag.*, vol. 14, no. 1, pp. 12–24 , 1995.

[18] P. Castia, A. Mencattinia, M. Salmeria, A. Anconab, F. Mangerib, M. L. Pepec and R. M. Rangayyand. "Estimation of the breast skin-line in mammograms using multidirectional Gabor filters", *Computers in Biology and Medicine*, vol. 43, no. 11, pp. 1870–1881 , 2013.

[19] J. Chakraborty, S. Mukhopadhyay, V. Singla, N. Khandelwal and P. Bhattacharyya. "Automatic detection of pectoral muscle using average gradient and shape based feature", *J. Dig. Imaging.*, vol. 25, no. 3, pp. 387–399, 2012.

[20] R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, R. A. Borges and A. F. Frère. "Automatic identification of the pectoral muscle in mammograms", *IEEE Trans. Med. Imag.*, vol. 23, no. 2, pp. 232–245, 2004.

[21] S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel and M. T. Rickard. "Automatic Pectoral Muscle Segmentation on Mediolateral Oblique View Mammograms", *IEEE Trans. Med. Imag.*, vol. 23, no. 9, pp. 1129–1140, 2004.

[22] N. Karssemeijer. "Automated classification of parenchymal patterns in mammograms", *Phys. Med. Biol.*, vol. 43, no. 2, pp. 365–378, 1998.

[23] C. Tsiotsios and M. Petrou. "On the choice of the parameters for anisotropic diffusion in image processing", *Pattern Recognition*, vol. 46, no. 5, pp. 1369–1381, 2013.

[24] P. Perona and J. Malik. "Scale space and edge detection using anisotropic diffusion", *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 629–639, 1990.

[25] N. Otsu. "A threshold selection method from gray-level histograms", *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.

[26] J. Suckling *et al.*. "The mammographic image analysis society digital mammogram database", *in Proc. Excerpta Med. Int. Congr. Ser.*, 1994, pp. 375–378.

[27] A. Rampun, L. Zheng, P. Malcolm, B. Tiddeman and R. Zwiggelaar. "Computer-aided detection of prostate cancer in T2-weighted MRI within the peripheral zone", *Phys. Med. Biol.*, vol. 61, no. 13, pp. 4796–4825, 2016.

[28] M. J. Black, G. Sapiro, D. Marimont and D. Heeger. "Robust anisotropic diffusion", *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 421–432, 1998.

[29] D. Mumford and J. Shah. "Optimal approximation by piecewise smooth functions and associated variational problems", *Commun. Pure Appl. Math*, vol. 42, pp. 577–685, 1989.

[30] K. K. Camilus, V. K. Govindan and P. S. Sathidevi. "Computer-aided identification of the pectoral muscle in digitized mammograms", *J. Dig. Imaging.*, vol. 23, no. 5, pp. 562–580, 2010.

[31] M. Awrangjeb and Guojun Lu. "Robust image corner detection based on the chord-to-point distance accumulation technique", *IEEE Trans. Multimedia*, vol. 10, no. 6, pp. 1059–1072, 2008.

[32] A. Oliver, X. Lladó, A. Torrent and J. Martí,. "One-shot segmentation of breast, pectoral muscle, and background in digitised mammograms", *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 912–916.

[33] J. Canny. "A Computational Approach to Edge Detection", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, No. 6, pp. 679–698, 1986.

[34] MatWorks, 'Regionprops', 2016. [Online]. Available: `https://uk.mathworks.com/help/images/ref/regionprops.html`. [Accessed:23-Sept-2016]

[35] MatWorks, 'Filtering and smoothing data', 2016. [Online]. Available: `https://uk.mathworks.com/help/curvefit/smoothing-data.html#bq_6ys3-8`. [Accessed:23-Sept-2016]

[36] S. D. Tzikopoulos, M. E. Mavroforakis, H. V. Georgiou, N. Dimitropoulos and S. Theodoridis."A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry", *Comput. Meth. Prog. Bio.*, vol. 102, no.1 , pp. 47–63, 2011.

[37] M. Wirth, D. Nikitenko and J. Lyon."Segmentation of the breast region in mammograms using a Rule-Based fuzzy reasoning algorithm", *Int. J. Graph. Vision Image Process.*, vol. 5, no.2 , pp. 45–54, 2005.

[38] K. S. Camilus, V. K. Govindan and P. S. Sathidevi."Pectoral muscle identification in mammograms", *Journal of Applied Medical Clinical Physics*, vol. 12, no.3, pp. 215–230, 2011.

[39] L. Liu, Q. Liu and W. Lu."Pectoral muscle detection in mammograms using local statistical features.", *J. Digit. Imaging*, vol. 25, no.5, pp. 633–641, 2014.

[40] V. Caselles, R. Kimmel and G. Sapiro."Geodesic active contours", *Int. J. Comput. Vision*, vol. 22, no.1, pp. 61–79, 1997.

[41] I. C. Moreira, I Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, J. S. Cardoso. INbreast: toward a full-field digital mammographic database. *Acad Radiol.* vol. 19, no. 2, pp. 236–428, 2011.

[42] M. A. G. Lopez, N. G. de Posada, D. C. Moura, R. R. Pollan, J. M. F. Valiente, C. S. Ortega, M. R. del Solar, G. D. Herrero, I. M. A. P. Ramos, J. P. Loureiro, T. C. Fernandes, B. M. F. de Arajo. BCDR: A BREAST CANCER DIGITAL REPOSITORY". ICEM15: 15th International Conference on Experimental Mechanics, FEUP-EURASEM-APAET, Porto/Portugal, 22-27 July 2012. ISBN: 978-972-8826-26-02.

[43] M. Ciecholewski. An edge-based active contour model using an inflation/deflation force with a damping coefficient. *Expert Systems with Applications*, vol. 44, pp. 22–36, 2016.

[44] L. Álvarez, L. Baumela, P. Henríquez and P. Márquez-Neila. "Morphological Snakes", *in Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2197–2202.

[45] J. V. Miller, D. E. Breen, and M. J. Wozny, Extracting geometric models through constraint minimization, Rensselaer Polytechnic Institute, Tech. Rep. No. 90024, 1990.

[46] J. V. Miller, D. E. Breen, W. E. Lorensen, R. M. OBara, and M. J. Wozny, Geometrically deformed models: A method to extract closed geometric models from volume data, Compur. Graphics, vol. 25, no. SPIE, vol. 1383, pp. 257–264, 1990.

[47] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, Member, IEEE, and John C. Gore. A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities With Application to MRI. *IEEE Trans. on Image Processing*, Vol. 20 (7), pp. 2007–2016, 2011.

[48] K. Zhang, H. Song, and L. Zhang. Active contours driven by local image fitting energy. *Pattern Recognition* Vol. 43 (4), pp. 1199–1206, 2010.